



US Army Corps
of Engineers®

Initial Development and Testing of a State-of-the-Art Method to Quantify Hydrologic Model Uncertainty

by Brian E. Skahill

PURPOSE: The purpose of this document is to report upon the initial independent implementation and testing of the Bayesian Markov Chain Monte Carlo (MCMC) algorithm Differential Evolution Markov Chain (DE-MC).

INTRODUCTION: To support the US Army Corps of Engineers' (USACE) use of risk-based analysis in flood damage reduction studies, Skahill (2012) identified the most promising state-of-the-art and practice-oriented approaches to robustly quantify hydrologic and hydraulic (H&H) model uncertainty. Skahill (2012) provides a path forward for related work activities, including software development, preparation of practice-oriented guidance documentation, and research and development directed at improving uncertainty analysis algorithm efficiency.

Bayesian Markov Chain Monte Carlo (MCMC), and in particular DiffereNtial Evolution Adaptive Metropolis (DREAM) (Vrugt et al. 2008a, 2009), and/or its basis, Differential Evolution Markov Chain (DE-MC) (ter Braak 2006), was selected by Skahill (2012) as the state-of-the-art method for estimating model parameter and predictive uncertainty. The intent of MCMC is to sample (upon completion of the burn-in period), via stochastic simulation, from the noted target equilibrium (i.e., posterior probability) distribution. The purpose of this document is to report upon the initial independent implementation and testing of the Bayesian MCMC algorithm DE-MC. This new DE-MC implementation differs notably from other MCMC implementations in that additional sampler burn-in (burn-in is the initial period when the MCMC sampler has not yet converged to its target equilibrium distribution) assessment heuristics were incorporated into the algorithm. These heuristics attempt to support a more robust assessment of sampler burn-in, rather than solely relying upon a quantitative sampler convergence diagnostic which can frequently prematurely misdiagnose convergence to the equilibrium target distribution. Two nontrivial test cases serve as a means to verify the independent DE-MC implementation and demonstrate related capabilities. The first test case involves a bimodal mixed normal target distribution and the second test case involves an application of the Sacramento Soil Moisture Accounting (SAC-SMA) hydrologic model. This technical note concludes with brief remarks regarding additional planned USACE-ERDC Bayesian MCMC research and development.

BAYESIAN MCMC DE-MC: Markov Chain Monte Carlo (MCMC) is a formal Bayesian approach for estimating the posterior probability distribution of specified adjustable model parameters. The idea behind MCMC is that while one wants to compute a probability density, $p(\theta|y)$, where θ and y represent the vector of adjustable model parameters and the observed data, respectively, there is the understanding that such an endeavor may be impracticable. Additionally, simply being able to generate a large random sample from the probability density would be equally sufficient as knowing its exact form. Hence, the problem then becomes one of effectively and efficiently generating a large number of random draws from $p(\theta|y)$. It was discovered that an

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE SEP 2013		2. REPORT TYPE		3. DATES COVERED 00-00-2013 to 00-00-2013	
4. TITLE AND SUBTITLE Initial Development and Testing of a State-of-the-Art Method to Quantify Hydrologic Model Uncertainty				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Engineer Research and Development, Center, Vicksburg, MS, 39180				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 15	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

efficient means to this end is to construct a Markov chain, a stochastic process of values that unfold in time, with the following properties: (1) the state space (set of possible values) for the Markov chain is the same as that for θ ; (2) the Markov chain is easy to simulate from; and (3) the Markov chain's equilibrium distribution is the desired probability density $p(\theta|y)$. By constructing such a Markov chain, one could then simply run it to equilibrium and subsequently sample from its stationary distribution. A Markov chain with the above mentioned properties can be constructed by choosing a symmetric proposal distribution and employing the Metropolis acceptance probability (Metropolis et al. 1953) to accept or reject candidate points.

MCMC simulation is more efficient than other Monte Carlo methods. The ability to sample from the posterior probability distribution for the specified adjustable model parameters, $p(\theta|y)$, provides one with the capacity to robustly address questions associated with the deployed modeled scenarios/alternatives from a probabilistic perspective. For example, a question such as "What is the probability a specific simulated state variable such as flow will be exceeded?", which of course has direct application not only to flood risk management and hydrologic design, but also to environmental and water quality analysis, among possible others, can be robustly answered via application of MCMC. Moreover, modeled sources of uncertainty; viz., input forcing, model parameter, and model structure, can all be encapsulated in θ in attempts to completely quantify model uncertainty (Vrugt et al. 2008a, b).

C. ter Braak (2006) introduced Differential Evolution Markov Chain (DE-MC), which combines the salient features of the global optimization method Differential Evolution (DE) (Storn and Price 1995, 1997) with Bayesian Markov Chain Monte Carlo. Multiple chains are run in parallel with DE-MC and learn from each other by way of jump proposals that are generated by taking the difference of two randomly selected chains from the current population. The probability of selecting the jump proposal is determined by using the Metropolis algorithm (Metropolis et al. 1953). The proposal vector for the simple but effective DE-MC algorithm is given as follows:

$$\mathbf{x}_p = \mathbf{x}_i + \gamma (\mathbf{x}_{R1} - \mathbf{x}_{R2}) + \mathbf{e} \quad (1)$$

where \mathbf{x}_p , \mathbf{x}_i , γ , \mathbf{x}_{R1} , \mathbf{x}_{R2} , and \mathbf{e} represent the proposal vector of dimension d , the i^{th} of the N chains which constitute the evolving population, a weighting factor, two unique vectors from the current population, excluding the i^{th} chain, randomly selected without replacement, and an error term sampled randomly from a symmetric distribution with small variance compared with that of the target. N and γ constitute the two parameters associated with the DE-MC method. C. ter Braak (2006) suggested $N = 2d$ or $3d$ for simple unimodal targets and $N = 10d$ to $20d$ for more complicated target distributions. Assuming the target distribution is multivariate normal, the optimal choice for γ is $2.38/2d$, and this value was observed to work well for the tests and examples considered by C. ter Braak (2006). Moreover, C. ter Braak (2006) demonstrated that adapting the DE-MC algorithm such that γ equals one every tenth generation mitigates against the potential of becoming trapped in a single mode within a multimodal distribution.

INDEPENDENT IMPLEMENTATION AND TESTING OF DE-MC: The author of this technical note wrote an independent implementation of the MCMC method DE-MC, as documented herein. Noteworthy elements of the independent implementation include the following:

1. Treatment of the initialization of the population.

2. Treatment of jump proposal dimensions that are outside of their bounds.
3. Use of the Gelman and Rubin GR diagnostic (Gelman and Rubin 1992), R , to assess convergence to the stationary target distribution.

Initialization of Population. The population to be evolved can be initialized either via uniform random sampling (URS) or Latin hypercube sampling (LHS). Based on limited experimentation to date, it is recommended that one employ LHS to initialize the population.

Treatment of Proposal Dimensions that are out of Bounds. For DE-MC, wherein jump proposals are defined by Equation 1, it is possible for one or more dimensions associated with the proposal of a given chain to be out of bounds. In that case, the specific proposal dimension which is out of bounds can either be (a) set equal to its appropriate boundary value, or alternatively it can be (b) set equal to its current value. Limited experience to date suggests recommending the use of the latter option (i.e., (b)) for this implementation issue.

Use of the Gelman and Rubin Convergence Diagnostic. A notable departure of the independent DE-MC implementation described herein from previous hydrologic modeling studies, which have employed Bayesian MCMC for optimization and inference, is how the Gelman and Rubin quantitative diagnostic, R , is employed to assess sampler convergence; viz., how R is used to determine if the burn-in period for the MCMC sampler is complete in that the chains have reached equilibrium and sampling is now from the target distribution. The Gelman and Rubin convergence diagnostic effectively measures the within and between variance of the chains, which are effectively the same when the chains have sufficiently mixed. A common specified value for R to diagnose sampler convergence is 1.2, and when the threshold value is reached, for each element of θ , it is often assumed that the MCMC sampler has converged to the target distribution (Vrugt et al. 2003; Vrugt et al. 2008a, 2008b; Vrugt et al. 2009). The interested reader is directed to Cowles and Carlin (1996), and references cited therein, for more detail and discussion pertaining to the Gelman and Rubin quantitative convergence diagnostic. Unfortunately, it is known that diagnostic has the tendency to prematurely assess MCMC sampler convergence to its stationary distribution (Kass et al. 1998). This characteristic for R was observed during initial testing and evaluation of the independent DE-MC implementation.

In attempts to ensure a more robust assessment of DE-MC sampler burn-in, a hybrid semi-automated approach was implemented, consistent with available guidance regarding practical application of Bayesian MCMC. Effectively, additional heuristics have been incorporated to assist with diagnosing sampler convergence. In particular, if the Gelman and Rubin convergence diagnostic indicates convergence, in that for each adjustable model parameter the value for R is less than a pre-specified threshold value, then diagnostic information associated with each chain and for each parameter is written to file for manual review. In addition, the MCMC algorithm is paused, the user is prompted that indicates sampler convergence and that data is available for evaluation to support determination as to whether or not the burn-in period is complete. If the user decides that the sampler has not converged, then the algorithm proceeds as previously, but with a reduced value for R , the new threshold value given by $(1+R_{prev})/2$. Alternatively, if the user decides that the DE-MC sampler has converged, then subsequent sampling during the post burn-in monitoring period proceeds unless it is determined that convergence was identified prematurely based on a comparison of the currently computed minimum RMSE fitness value versus the minimum root-mean-square error (RMSE) fitness value computed at what was

previously thought to be the completion of the burn-in phase for the sampler. In this case, burn-in sampling recommences with a new value for R as mentioned previously. The threshold comparison value of the noted minimum RMSEs, which suggests that burn-in was previously assessed prematurely, is a specified user input. This additional heuristic incorporated into the algorithm to automatically self-correct during post burn-in monitoring is relatively simplistic and requires further evaluation and testing. Testing to date, has demonstrated improved assessment of DE-MC MCMC sampler convergence.

Verification and Testing. Verification of the independent DE-MC implementation involved two documented tests: one, a non-trivial known target distribution (ter Braak 2006), and the other involving calibration of a Sacramento Soil Moisture Accounting (SAC-SMA) hydrologic model deployment (Vrugt et al. 2009).

Bimodal mixed normal target distribution

Numerical tests to validate the noted independent DE-MC implementation were performed with a known bimodal target composed of the sum of two normal distributions, which is documented to be a difficult problem for MCMC simulation in that the distance between the two modes makes jump proposals from one mode to the other problematic (ter Braak 2006; Vrugt et al. 2009). In particular, numerical experiments were performed for the specified known bimodal target:

$$\pi(\mathbf{x}) = \frac{1}{3}N_d(-5, I_d) + \frac{2}{3}N_d(5, I_d) \quad (2)$$

where π , \mathbf{x} , N_d , $\mathbf{5}$, and I_d represent the target distribution, a given chain in the population, a multivariate normal distribution of dimension d , a vector consisting of a 5 in each of the d -dimensions, and a d -dimensional identity matrix, respectively.

Consistent with ter Braak (2006), all of the tests with the known bimodal target specified in Equation 2, involved (1) a predetermined burn-in of 1000 generations, (2) two different initial distributions; viz., a narrow initial distribution given by $N(\mathbf{0}, I_d)$ and a broad initial distribution given by $N(2.5, 25 I_d)$, (3) two values specified for d and N ; viz., 5 and 10, and 100 and 1000, respectively, and (4) γ set equal to 2.38/2. In each case, 100 independent trials were performed and uniform random sampling was employed to initialize the population. The subsequent monitoring period consisted of 1000 generations. Table 1 summarizes the trial runs. Figures 1 – 7 provide an additional summary of the numerical experiments. For $d = 10$ and with the default value specified for γ , it was observed, as is clearly evident upon inspection of Figures 4 and 6, that DE-MC had difficulty converging to the known target distribution. However, activation of the adaption functionality for γ , previously mentioned and briefly discussed above, remedied the observed problem, as indicated in Table 1 and also in Figures 5 and 7. For $d = 10$ and $N = 1000$, additional numerical experiments, each consisting of 100 independent trials, were also performed wherein γ was effectively set equal to one every fifth, fourth, and third generation, in attempts to examine how the update frequency would impact the acceptance rate and also the RMSE for the estimate of the expected value (1.666). Table 2 summarizes the results from these additional numerical experiments.

Table 1. Summary of numerical experiments with independent DE-MC implementation with the known bimodal target distribution (If yes for adaption, then, as indicated in the text of the note, γ is effectively set equal to one every tenth generation; RMSE for estimate of the expected value, which is 1.666). Results based on the 100 independent trials.

d	N	Initial Population	Adaption	Acceptance Rate	RMSE
5	100	Narrow	No	0.163	0.37
5	1000	Narrow	No	0.157	0.112
5	100	Broad	No	0.166	0.352
5	1000	Broad	No	0.159	0.108
10	1000	Narrow	No	0.13	1.609
10	1000	Narrow	Yes	0.129	0.124
10	1000	Broad	No	0.199	2.054
10	1000	Broad	Yes	0.145	0.23

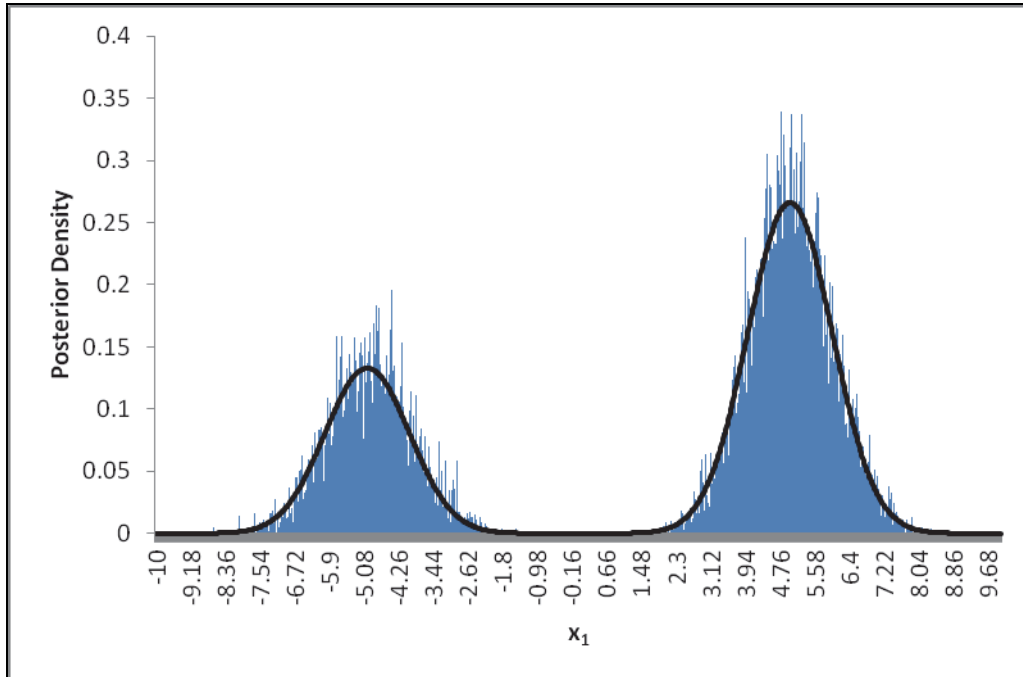


Figure 1. Estimated marginal posterior probability distribution of x_1 from trial 33, based on using $d = 5$, $N = 100$, and the narrow initial population. The solid black line depicts the true bimodal target distribution.

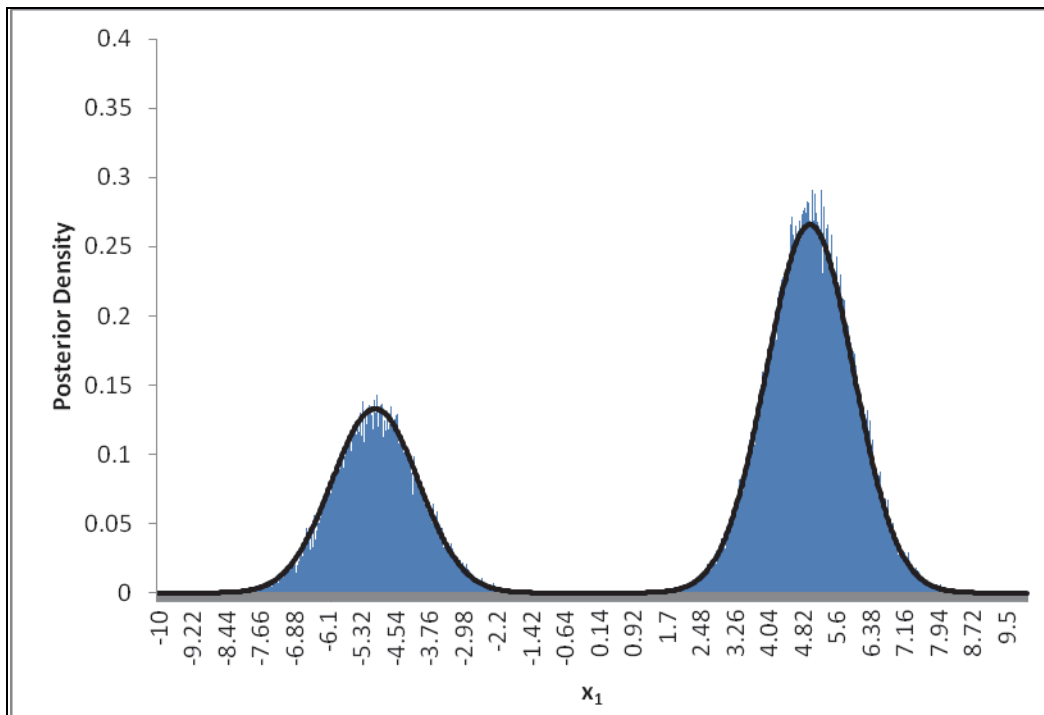


Figure 2. Estimated marginal posterior probability distribution of x_1 from trial 2, based on using $d = 5$, $N = 1000$, and the narrow initial population. The solid black line depicts the true bimodal target distribution.

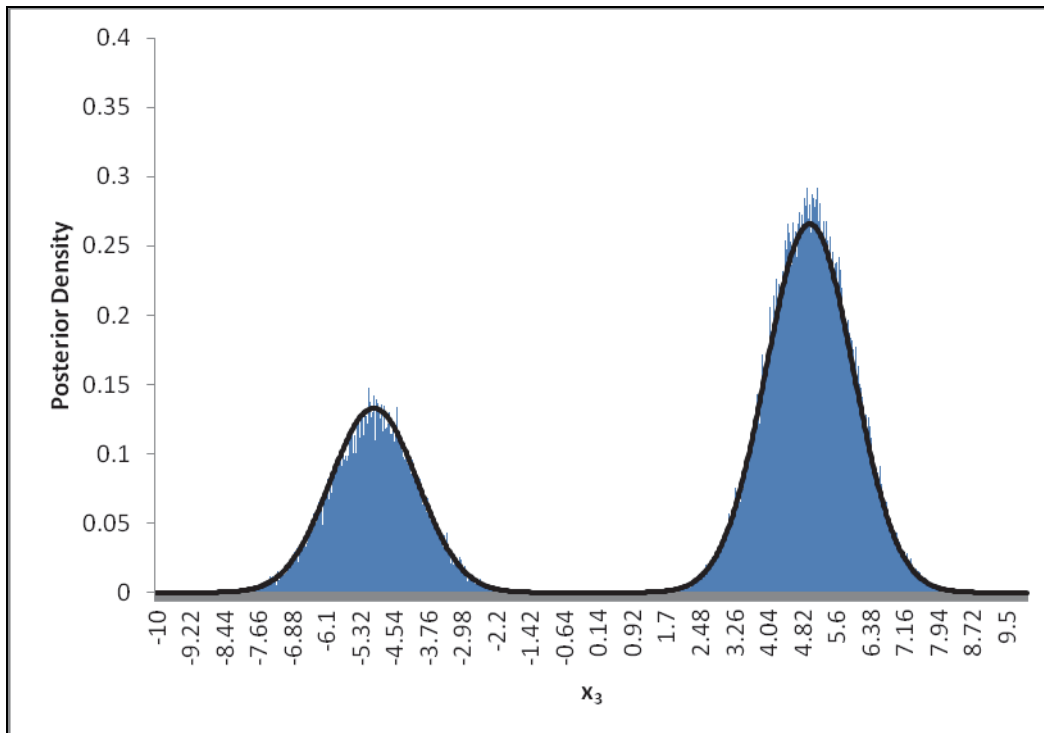


Figure 3. Estimated marginal posterior probability distribution of x_3 from trial 2, based on using $d = 5$, $N = 1000$, and the broad initial population. The solid black line depicts the true bimodal target distribution.

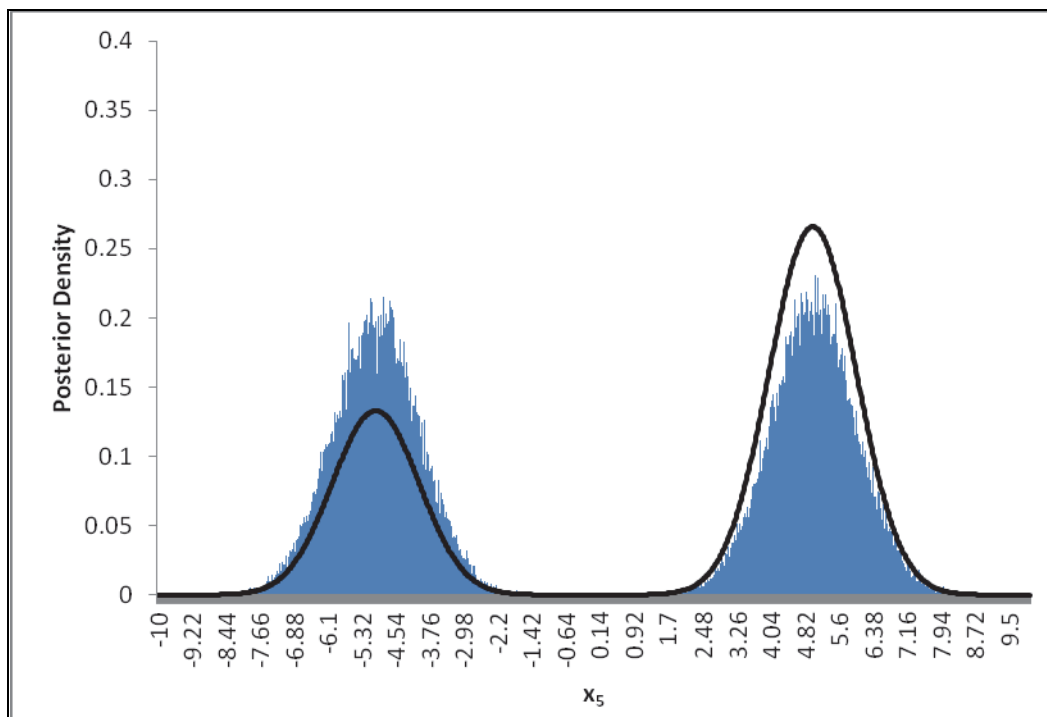


Figure 4. Estimated marginal posterior probability distribution of x_5 from trial 3, based on using $d = 10$, $N = 1000$, the narrow initial population, and not adapting γ . The solid black line depicts the true bimodal target distribution.

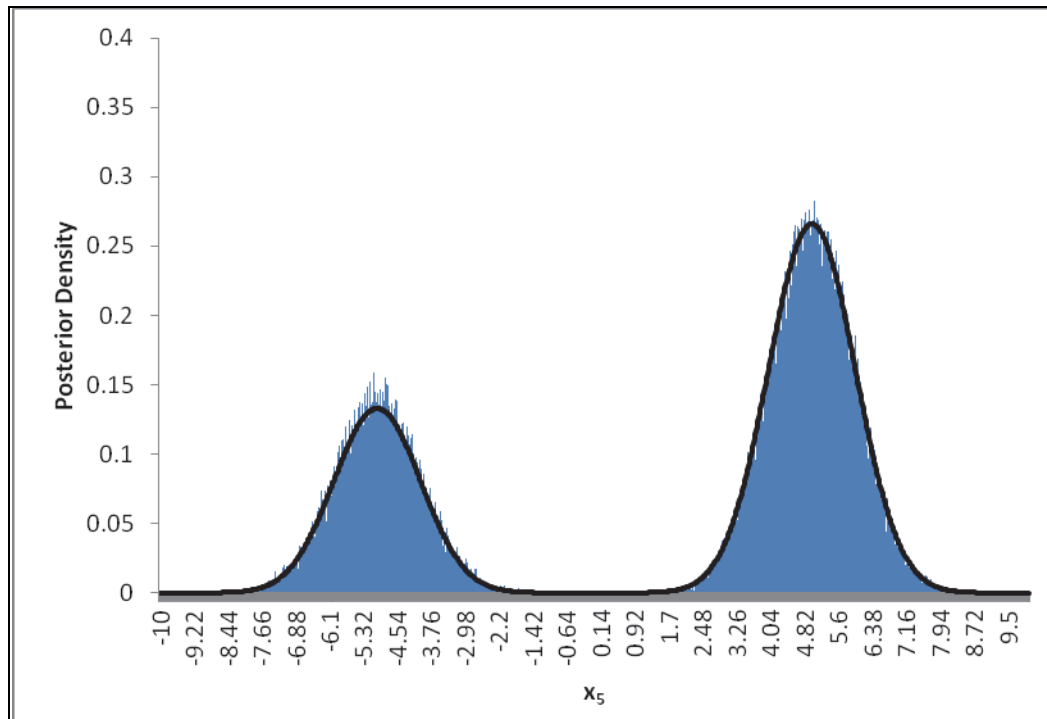


Figure 5. Estimated marginal posterior probability distribution of x_5 from trial 5, based on using $d = 10$, $N = 1000$, the narrow initial population, and setting γ equal to one effectively every tenth generation. The solid black line depicts the true bimodal target distribution.

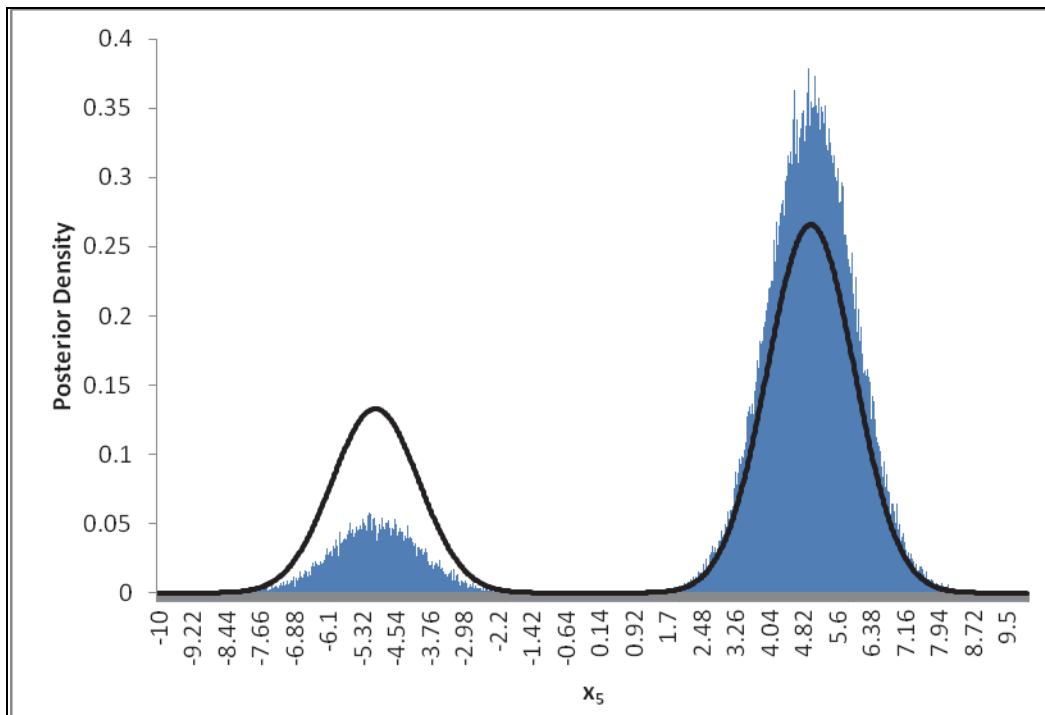


Figure 6. Estimated marginal posterior probability distribution of x_5 from trial 4, based on using $d = 10$, $N = 1000$, the broad initial population, and not adapting γ . The solid black line depicts the true bimodal target distribution.

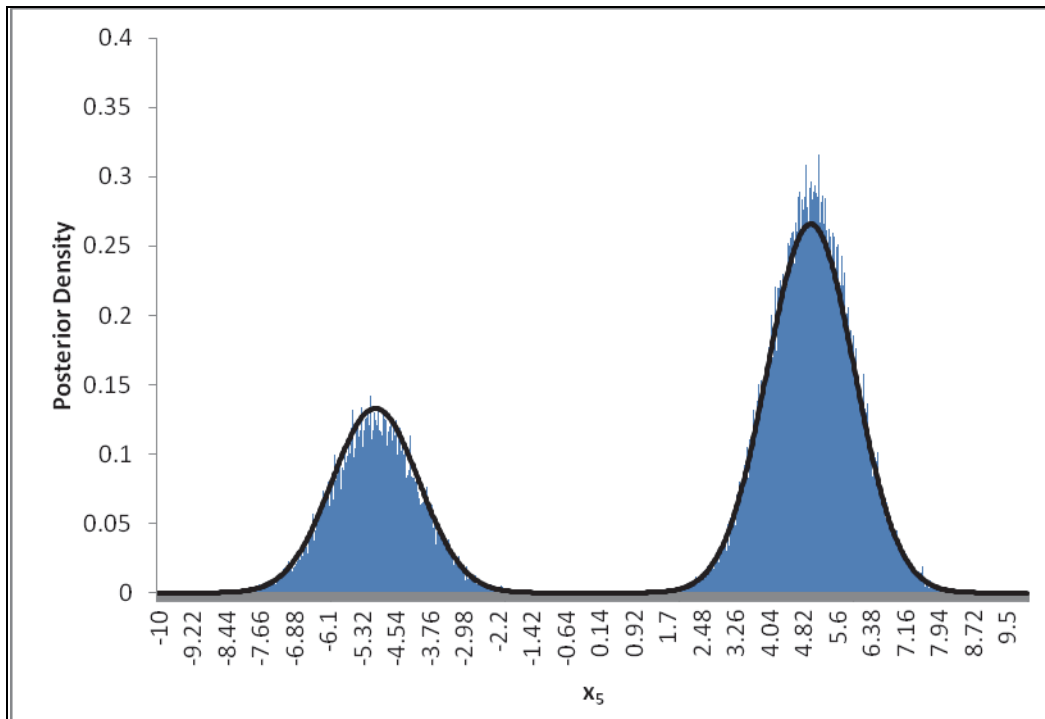


Figure 7. Estimated marginal posterior probability distribution of x_5 from trial 5, based on using $d = 10$, $N = 1000$, the broad initial population, and setting γ equal to one effectively every tenth generation. The solid black line depicts the true bimodal target distribution.

Table 2. Summary of the impact of the update frequency for γ.					
d	N	Initial Population	Adaption Update Frequency	Acceptance Rate	RMSE
10	1000	Narrow	0.1	0.129	0.124
10	1000	Narrow	0.2	0.120	0.091
10	1000	Narrow	0.25	0.115	0.085
10	1000	Narrow	0.333	0.107	0.086
10	1000	Broad	0.1	0.145	0.230
10	1000	Broad	0.2	0.127	0.109
10	1000	Broad	0.25	0.119	0.093
10	1000	Broad	0.333	0.109	0.085

SAC-SMA hydrology model

Vrugt et al. (2009) described the DREAM MCMC sampler and compared it with other MCMC samplers via several case studies, one which involved inferring the posterior probability distribution for thirteen Sacramento Soil Moisture Accounting (SAC-SMA) hydrologic model parameters for a SAC-SMA deployment to the 1944 km² Leaf River watershed near Collins, MS using two years of mean daily flow data. The SAC-SMA hydrologic model is used by the National Weather Service (NWS) for flood forecasting throughout the United States. While it has sixteen parameters that need to be specified by the user, consistent with previous work (Vrugt et al. 2003 and references cited therein), thirteen were specified as adjustable. The prior uncertainty ranges of the specified adjustable SAC-SMA hydrologic model parameters are defined in Table 3.

Table 3. SAC-SMA model parameters, including their prior, and units.			
Parameter	Minimum	Maximum	Unit
UZWIM	1	150	[mm]
UZFWM	1	150	[mm]
UZK	0.1	0.5	day-1
PCTIM	0	0.1	[-]
ADIMP	0	0.4	[-]
ZPERC	1	250	[-]
REXP	1	5	[-]
LZWIM	1	500	[mm]
LZFSM	1	1000	[mm]
LZFPM	1	1000	[mm]
LZSK	0.01	0.25	day-1
LZPK	0.0001	0.025	day-1
PFREE	0	0.6	[-]

The reader is referred to Vrugt et al. 2003, and references cited therein, for comprehensive discussions regarding the SAC-SMA hydrologic model, the Leaf River watershed and its related hydrologic data (viz., mean areal precipitation (mm/day), potential evapotranspiration (mm/day), and streamflow (m³/s)) which was used to support the inference of the posterior distribution of the SAC-SMA specified adjustable model parameters.

Consistent with Vrugt et al. (2009), for the series of implementation validation and evaluation tests employed with this “real world” case study application, squared deviation likelihood was utilized as follows:

$$\pi(x) = \left[\sum_{i=1}^T (S_i(x) - \hat{S}_i)^2 \right]^{-\frac{1}{2}T} \quad (3)$$

where \hat{S}_i , S_i , and T represent the streamflow observations, their model simulated counterparts, and the number of data comparisons used to evaluate the likelihood (viz. 731). With each SAC-SMA forward model call, approximately two months of simulation occurs prior to the window for performing comparisons of model simulated flows with their observed counterparts.

For purposes of illustration, Figure 8 displays results associated with five unique realizations of the DE-MC method to calibrate the SAC-SMA Leaf River hydrologic model. In particular, it includes five plots of the RMSE computed each evolution. For each of the five DE-MC simulation runs, $N=26$ and γ was effectively set equal to one every fifth generation. The observed differences in Figure 8 across the five RMSE plots is a function of the initial seed specified to the DE-MC implementation as all of the other input settings remained constant. The five plots in Figure 8 not only underscore the well known documented difficulty of calibrating the SAC-SMA conceptual hydrologic model (Gupta et al. 2003; Duan et al. 1992), but also the variability in the performance of the DE-MC method. One can clearly observe six unique optima upon inspection of the five RMSE plots in Figure 8. Moreover, one can also observe from the five plots that the number of forward model calls necessary to achieve burn-in for the DE-MC sampler can exhibit a fair degree of variability from one simulation run to another. With the independent DE-MC implementation, a minimum RMSE value of 13.25 was achieved, as was also reported upon in Vrugt et al. (2009). Figure 9 is a trace plot of the specified SAC-SMA adjustable model parameter UZTWM associated with the first chain for one of the five DE-MC simulations. Figure 10 is a plot of the 95 percent predictive uncertainty bounds associated with one of the five DE-MC simulation runs. The computed uncertainty bounds are solely a function of the model parameter uncertainty obtained via sampling from the posterior probability density during the post burn-in monitoring period which was arbitrarily set to be equal in length to the number of simulations required to achieve DE-MC sampler burn-in.

DISCUSSION AND CONCLUSIONS: This technical note briefly describes the Differential Evolution Markov Chain MCMC sampler, which was identified in Skahill (2012) to be a state-of-the-art methodology for quantifying hydrologic model parameter and predictive uncertainty in support of risk-based hydrologic design, and evaluation of project alternatives. This document also describes some of the salient features associated with an independent implementation of the DE-MC method; viz., approaches to initialization of the population to be evolved, treatment of jump proposal dimensions that are out of bounds, and a hybrid, heuristic, semi-automated approach for assessing convergence of the DE-MC sampler. In addition, this technical note summarizes two case study applications of the independent DE-MC implementation.

The first case study involved a bimodal mixed normal target distribution, a non-trivial verification problem in that in the past been demonstrated to be difficult to resolve for

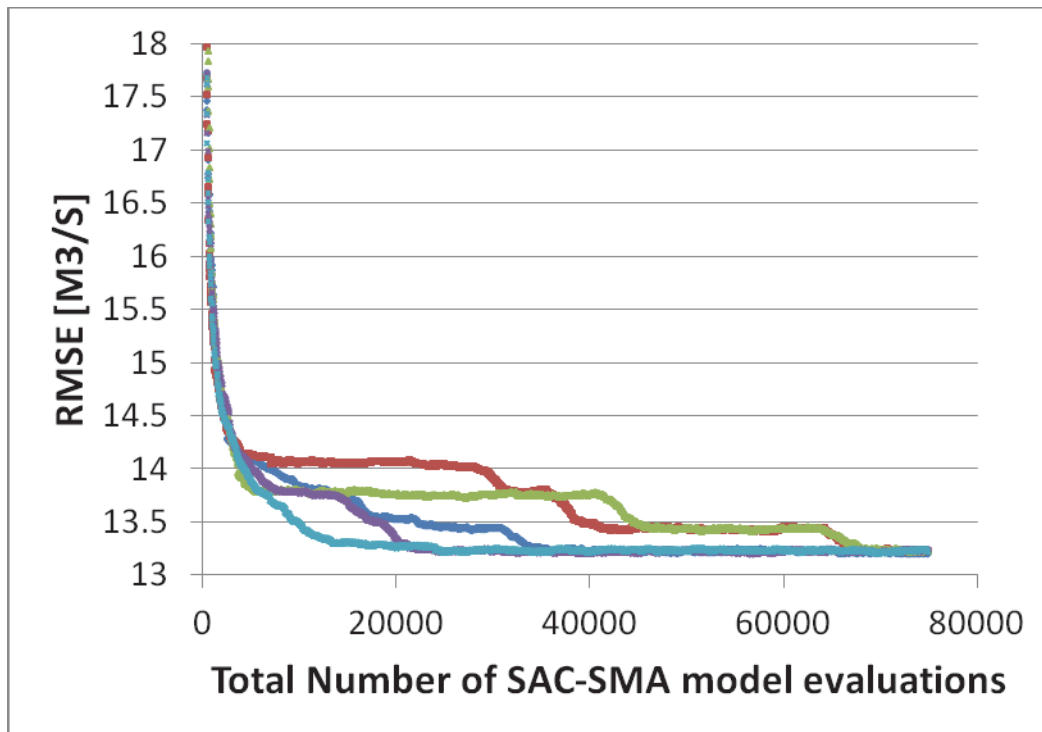


Figure 8. A plot of mean RMSE for each generation for five unique DE-MC runs.

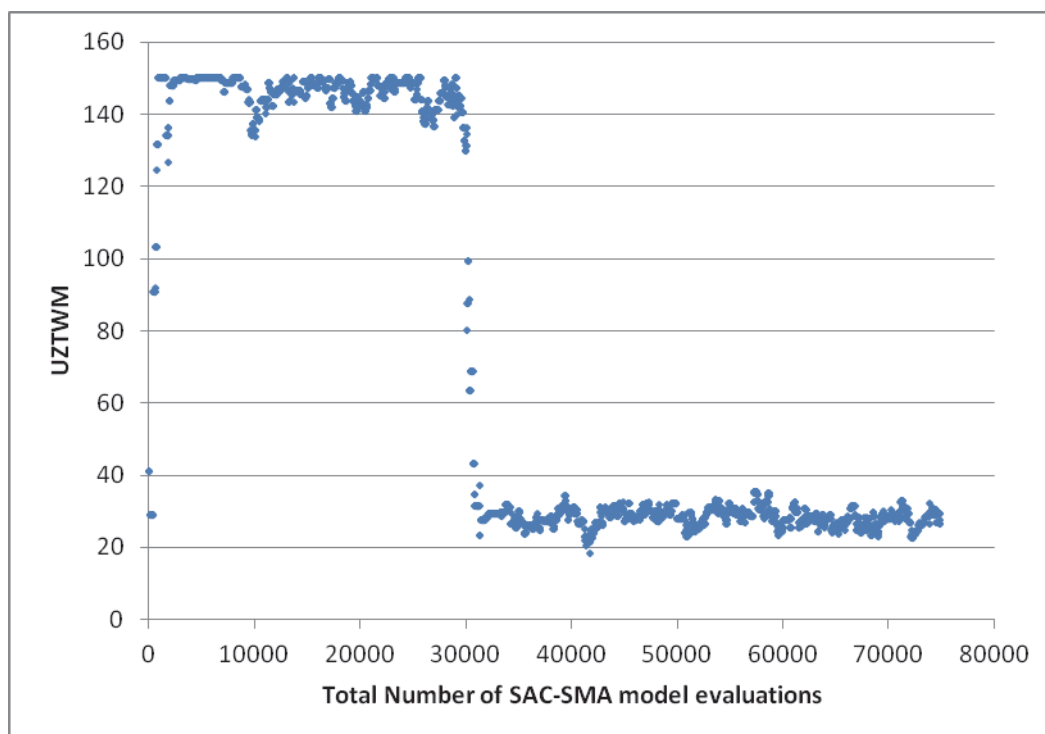


Figure 9. A plot of the evolution of the first chain for the SAC-SMA parameter UZTWM.

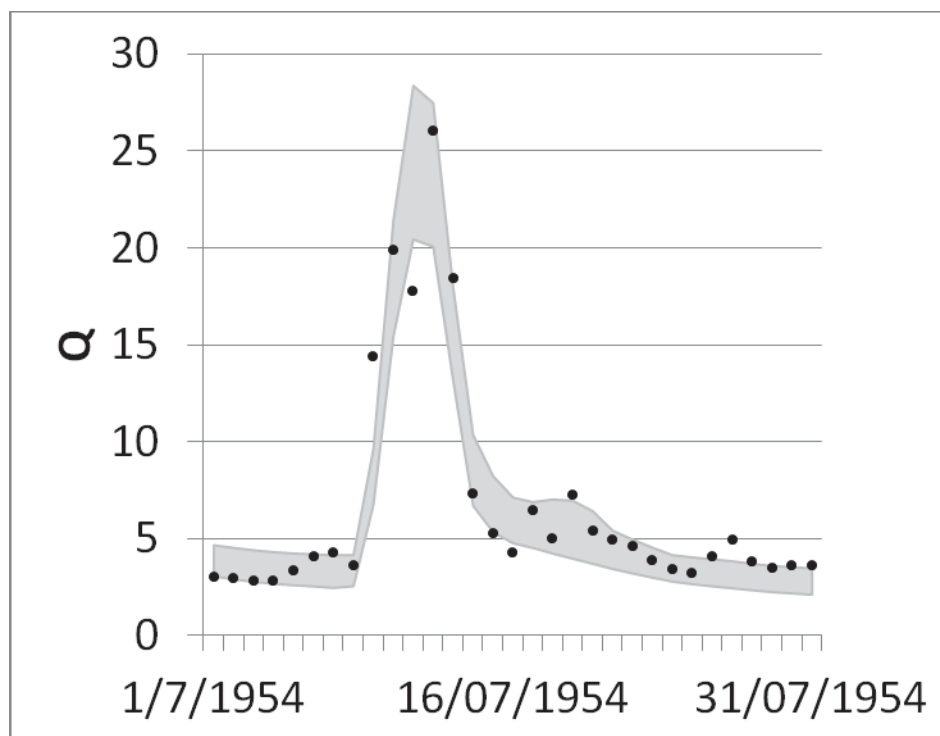


Figure 10. 95 percent predictive uncertainty envelope which is solely a function of model parameter uncertainty.

classical random walk metropolis MCMC samplers. However, systematic adaption of one of the two DE-MC algorithm input parameters ensured convergence to the bimodal target distribution. The second case study involved application of the independent implementation of the DE-MC method to calibrate a Sacramento Soil Moisture Accounting hydrology model deployment using two years of mean daily flow data. This validation and evaluation case study demonstrated that the DE-MC method is able to handle complex response surfaces and find the documented minimum. And that by sampling from the equilibrium posterior probability distribution for the specified adjustable model parameters one can then generate model predictive uncertainty bounds which can support risk-based analyses. Although, in the second case study example it was also illustrated that the DE-MC method has the capacity to exhibit a fair degree of random variability, when applied, in terms of the number of forward model calls required to achieve burn-in for the MCMC sampler.

The work documented herein is just the beginning with respect to research and development directed to the development of a state-of-the-art robust methodology for computing model parameter and predictive uncertainty. Opportunities for future DE-MC related study and exploration include:

1. Refine the current hybrid, heuristic, semi-automated approach for assessing DE-MC convergence by addressing the question “Can MCMC burn-in be completely automated and remain reliable?”.
2. Ensure implementation of the DE-MC algorithm is as efficient as possible and reduce the observed variability in DE-MC algorithm performance.
 - a. Would dynamic adjustment of the DE-MC parameter γ improve algorithm performance?

- b. Explore applying those aspects of the DREAM MCMC algorithm (Vrugt et al. 2008a, 2008b, 2009) which do not complicate its implementation and practical usage, such as dynamic chain removal.
3. Bayesian MCMC provides one with a robust means for assessing model uncertainty; however, it is computationally costly. Regardless of the efficiencies derived, port the final independent Bayesian MCMC implementation to the HPC framework.
4. Integrate the independent DE-MC implementation into the USACE-ERDC GSSHA hydrology model (Downer and Ogden 2003a, b) as an alternate simulation mode.
5. Enhance the DE-MC GSSHA alternate simulation mode to also account for precipitation forcing uncertainty via specification of event multipliers similar to the approach described in Vrugt et al. (2008b).
6. Modify the independent DE-MC implementation to account for model structural uncertainty via specification of a likelihood measure that accommodates for the treatment of serial correlation of the residuals. Explore alternative means to quantify model structural uncertainty.
7. Incorporate into the independent DE-MC implementation automated means for computing the necessary length of the post burn-in monitoring period.
8. Provide a more complete treatment and discussion of the various options/features of the independent implementation of the DE-MC method to provide potential users with a clear picture regarding recommended usage of various options such as use of URS versus LHS for initializing the population to be evolved or the treatment of jump proposal dimensions that are out of bounds.
9. Clearly demonstrate the capacity of the Gelman and Rubin diagnostic to prematurely assess MCMC sampler convergence.
10. Perform comprehensive case studies that demonstrate use of the state-of-the-art model uncertainty analysis method to support practical risk-based hydrologic and environmental application settings.

ADDITIONAL INFORMATION: For additional information, contact Dr. Brian E. Skahill at Brian.E.Skahill@usace.army.mil. This CHETN should be cited as follows:

Skahill, B. E. (2013). *Initial development and testing of a state-of-the-art method to quantify hydrologic model uncertainty*. ERDC/CHL CHETN-VIII-7. Vicksburg, MS: US Army Engineer Research and Development Center. An electronic copy of this CHETN is available from <http://chl.erdcl.usace.army.mil/chetn>.

REFERENCES

- Cowles, M. K. and B. P. Carlin. 1996. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, 91(434): 883-904.
- Downer, C. W. and F. L. Ogden. 2003a. GSSHA user's manual: Gridded surface subsurface hydrologic analysis, version 1.43 for WMS 6.1. Technical Report. Vicksburg, MS: US Army Engineer Research and Development Center.
- Downer, C. W. and F. L. Ogden. 2003b. Prediction of runoff and soil moistures at the watershed scale: Effects of model complexity and parameter assignment. *Water Resources Research*, 39(3): 1045.

- Duan, Q. S., S. Sorooshian, and V. K. Gupta. 1992. Effective and efficient global optimization for conceptual rainfall runoff models. *Water Resour. Res.* 28 (4): 1015–1031.
- Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.*, 7, 457–472.
- Gupta, H. V., Sorooshian, S., Hogue, T.S., Boyle, D. P. 2003. Advances in automatic calibration of watershed models. In: Duan, Q., Gupta, H., Sorooshian, S., Rousseau, A., Turcotte, R. (Eds.), *Water Science and Application Series 6*: 197–211.
- Kass, R. E., Carlin, B. P., Gelman, A., and R.M. Neal. 1998. Markov Chain Monte Carlo in Practice: A Roundtable Discussion. *The American Statistician*, 52(2): 93-100.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J Chem Phys*, 21: 1087-1092.
- Skahill, B. E. 2012. Practice Driven and State-of-the-art Methods to Quantify Hydrologic Model Uncertainty. Technical Note. Vicksburg, MS: US Army Engineer Research and Development Center.
- Storn, R. and K. Price. 1995. Differential Evolution — a simple and efficient adaptive scheme for global optimization over continuous spaces. *International Computer Science Institute, Berkeley*, TR-95-012.
- Storn, R. and K. Price. 1997. Differential Evolution — a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimisation*, 11: 341–359.
- ter Braak, C. J. F. 2006. A Markov Chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces. *Stat. Comput.* 16: 239-249.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian. 2003. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.*, 39(8): 1201, doi:10.1029/2002WR001642.
- Vrugt, J. A., C. J. F. ter Braak, M. P. Clark, J. M. Hyman, and B.A. Robinson. 2008a. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research*, Vol. 44, W00B09, doi:10.1029/2007WR006720.
- Vrugt, J. A., C. J. F. ter Braak, H. V. Gupta, and B. A. Robinson. 2008b. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environ. Res. Risk Assess.*
- Vrugt, J. A., C. J. F. ter Braak, C.G.H. Diks, B.A. Robinson, J.M. Hyman, and D. Higdon. 2009. Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. *International Journal of Nonlinear Sciences & Numerical Simulation*, 10(3): 271-288.

NOTE: The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products.